



## Document Summary

---



[Preview Claims](#)

[Preview Full Text](#)

[Preview Full Image](#)

Email Link: 

**Document ID:** J P 07-249042 A2

**Title:** DOCUMENT PROCESSOR

**Assignee:** FUJI XEROX CO LTD

**Inventor:** YAMASHITA AKIO

**US Class:**

**Int'l Class:** G06F 17/27 A

**Issue Date:** 09/26/1995

**Filing Date:** 03/08/1994

### Abstract:

**PURPOSE:** To obtain a document processor which detects places to be proofread and elaborated in one specific text and also provides the proofreading and elaboration information capable of curvyng out required confirmation and correction on an editing device equipped with no function for proofreading and elaboration.

**CONSTITUTION:** The document processor is equipped with a text storage means (text storage part 2) which stores the text to be proofread and elaborated, a selection means (appearing character storage part 5 and specification part 5A attached thereto, and judgement part 6A and selection part 6B of output processing part 6) which selects characters undetected in the text as characters for marking by referring to the text to be proofread and elaborated in the text storage means, a detection processing means (text detection processing part 3) which detects candidate places to be proofread and elaborated from the text to be proofread and elaborated in the text storage means, and an insertion means (processing result storage part 4 and insertion part 6C of output processing part 6) which inserts the characters for marking selected by the selection means at the candidate places for proofreading and elaboration detected by the detection processing means.

(C)1995,JPO

---

Copyright © 1993-2000 Aurigin Systems, Inc.  
[Legal Notices](#)

(19)日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11)特許出願公開番号

特開平7-249042

(43)公開日 平成7年(1995)9月26日

(51)Int.Cl.<sup>6</sup>  
G 0 6 F 17/27

識別記号 庁内整理番号  
9288-5L

F I  
G 0 6 F 15/ 20 5 5 0 Z

技術表示箇所

審査請求 未請求 請求項の数2 F D (全 14 頁)

(21)出願番号 特願平6-62123

(22)出願日 平成6年(1994)3月8日

(71)出願人 000005496

富士ゼロックス株式会社  
東京都港区赤坂三丁目3番5号

(72)発明者 山下 明男

神奈川県川崎市高津区坂戸3丁目2番1号  
K S P R&Dビジネスパークビル 富士ゼロックス株式会社内

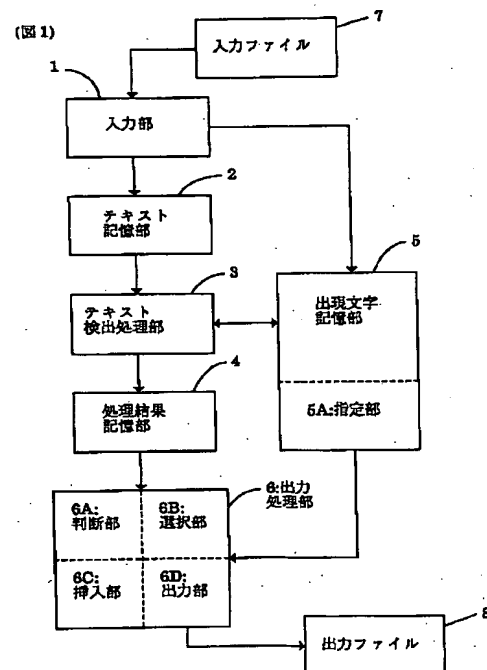
(74)代理人 弁理士 岩上 昇一 (外2名)

(54)【発明の名称】 文書処理装置

(57)【要約】

【目的】ある所定のテキストより校正・推敲すべき箇所を検出するための文書処理装置であって、校正・推敲の機能を備えていない編集装置において所期の確認・修正が可能にされる校正・推敲情報を提供できるものを提供すること。

【構成】校正推敲対象のテキストを記憶するテキスト記憶手段(テキスト記憶部2)、前記テキスト記憶手段に記憶された校正推敲対象のテキストを参照して、該テキスト中で検出されなかった文字をマーキング用の文字として選択する選択手段(出現文字記憶部5とそれに付属の指定部5A、出力処理部6における判断部6Aおよび選択部6B)、前記テキスト記憶手段に記憶された校正推敲対象のテキストから校正推敲候補の箇所を検出する検出処理手段(テキスト検出処理部3)、および、前記検出処理手段により検出された前記校正推敲候補の箇所に前記選択手段により選択されたマーキング用の文字を挿入する挿入手段(処理結果記憶部4および出力処理部6における挿入部6C)を具備して構成された文書処理装置。



#### 【特許請求の範囲】

【請求項1】校正推敲対象のテキストを記憶するテキスト記憶手段と、

前記記憶手段に記憶された校正推敲対象のテキストを参照して、該テキスト中で出現が検出されなかった文字をマーキング用の文字として選択する選択手段と、

前記テキスト記憶手段に記憶された校正推敲対象のテキストから校正推敲候補の箇所を検出する検出処理手段と、

前記検出処理手段により検出された前記校正推敲候補の箇所に前記選択手段により選択されたマーキング用の文字を挿入する挿入手段と、

を具備している文書処理装置。

【請求項2】マーキング用の文字の出現状態を記憶する出現文字記憶手段と、

前記出現文字記憶手段に出現状態を記憶すべきマーキング用の文字を指定する指定手段とを更に備え、

前記選択手段は、前記テキスト記憶手段に記憶されたテキスト中でのマーキング用の文字の出現状態を前記出現文字記憶手段に記憶するとともに、前記出現文字記憶手段に記憶したマーキング用の文字の出現状態に基づき目標のマーキング用の文字を選択することを特徴とする、前記請求項1に記載の文書処理装置。

#### 【発明の詳細な説明】

##### 【0001】

【産業上の利用分野】この発明は文書処理装置に関するものであり、特に、ある所定のテキストより校正・推敲すべき箇所を検出するための装置に関するものであって、校正・推敲の機能を備えていない編集装置において所期の確認・修正が可能にされる校正・推敲情報を提供することができる前記の文書処理装置に関するものである。

##### 【0002】

【従来の技術】近年、ワードプロセッサ等の文書処理装置の出現により、このような装置を用いて日本語文章を作成することが多くなってきている。しかるに、このようなことが多くなるにつれて、手書きのときには生じなかったような誤字・脱字の発生が起きやすくなり、その結果として、当該誤字・脱字の訂正がなされないままでの文書が作成されてしまうことが多くなってきた。また、複数人が分担して単一の文書を作成することも容易になってきているが、文章の表記法に個人差があることを放置していることから、その表記の仕方に統一性がない文章が作成されてしまうこともある。これに対して、ある所定の文書（テキスト）について、誤字・脱字や表記の不統一があるものを検出し、その確認・修正を助けるようにした校正・推敲支援システムの開発がなされている。従来は、ある所定の文書から検出された校正・推敲の箇所について、当該文書の別の箇所と区別できるマーキング操作を施し、前記検出された箇所が容易に識別

できるようにされていた。例えば特開平1-303564号公報「文書作成支援装置」によれば、装置によって指摘された文書の該当箇所について、校正内容の種類に応じたやり方で、所定の画面上で強調して表示するようにされている。また、特開平3-9465号公報「文章作成装置」によれば、校正・推敲の対象となる語句以外の語句を不可視状態または略不可視状態で表示させて、前記対象となる語句とそれ以外の語句とを区別して表示するようにされている。そして、このようなマーキングを手がかりにすることにより、装置の利用者は、対象とする文書における検出箇所を確認して、所期の修正が必要であったら対応の修正操作を施すようにされている。なお、前記の確認・修正がなされた後では、前記のマーキングは消去されることになる。なお、ある所定の文法処理に基づく校正・推敲支援機能はその処理が複雑になるものであることから、高性能の専用機に搭載されるのが通常であり、一般的な普及型のワードプロセッサやエディタ等に簡単に搭載できるものではなかった。そこで、通常の対応の仕方としては、前記一般的な普及型のワードプロセッサで作成した文書を校正・推敲支援機能付きの高級専用機側に移送し、この高級専用機側において校正・推敲情報の検出や所要の確認・修正作業をしたり、または、前記の高級専用機側において検出した校正・推敲情報を一旦印刷して前記ワードプロセッサ側に持ち帰り、この印刷の結果を目視で追いつながり確認して、これに対応する修正をワードプロセッサ側で行っていた。

##### 【0003】

【発明が解決しようとする課題】従来の校正・推敲支援システムにおいては、文書の所要箇所に「アンダーライン」、[シェード]、[色付け]等を施して検出すべき箇所を容易に認識できるようにされているが、マーキングの仕方や確認・修正のやり方が使用される装置に依存する性質のものであるために、当該使用される装置から外れると、該当の校正・推敲支援機能を直接的に利用することができないという問題点があった。これに対する一つの方策としては、前記の検出箇所を「◆」のような特殊な記号で囲むようにしておき、一般的なワードプロセッサの探索機能を用いてこの特殊記号を捜し出し、所要の確認・修正を支援することが考えられる。しかしながら、このような特殊記号を用いて検出箇所を囲むような対応の仕方では、当該特殊記号が校正推敲対象の文書に含まれている可能性があることから、マーキング用の特殊記号が校正推敲対象の文書に含まれていたかどうか、実際にマーキングのために用いられたものかどうかを確認する必要がある。また、前記確認・修正が終了した後でマーキング用の特殊記号を削除するようにされているが、このときにいわゆる「一括削除」を施すと、本来残すべき校正推敲対象の文書内のものまで削除してしまう可能性があり、削除に際しても一定の確認をするこ

とが必要である。

【0004】この発明は、上記された従来例や検出箇所を特殊記号で囲むやり方に存在する問題点に鑑みてなされたものであり、例えば、一般的なワードプロセッサから所要の校正・推敲支援を指定することが可能であって、検出箇所をマーキング用の特殊記号で示すことが可能にされた文書処理装置を提供することを目的とするものである。

#### 【0005】

【課題を解決するための手段】この発明に係る文書処理装置は、校正推敲対象のテキストを記憶するテキスト記憶手段と、前記記憶手段に記憶された校正推敲対象のテキストを参照して、該テキスト中で検出されなかった文字をマーキング用の文字として選択する選択手段と、前記テキスト記憶手段に記憶された校正推敲対象のテキストから校正推敲候補の箇所を検出する検出処理手段と、前記検出処理手段により検出された前記校正推敲候補の箇所に前記選択手段により選択されたマーキング用の文字を挿入する挿入手段とを具備していることを特徴とするものである。また、この発明に係る文書処理装置は、マーキング用の文字の出現状態を記憶する出現文字記憶手段と、前記出現文字記憶手段に出現状態を記憶すべきマーキング用の文字を指定する指定手段とを更に備えていることを特徴とするものである。

#### 【0006】

【作用】この発明によれば、テキスト記憶手段に記憶されている校正推敲対象のテキストから校正・推敲の候補の箇所（例えば、対象のテキストを構成する文章における誤字・脱字の存在箇所に関する情報）が、テキスト検出処理手段によって検出される。検出処理手段により検出された校正・推敲の候補の箇所には、選択手段により選択されたマーキング用の文字が挿入手段により挿入される。ここでのマーキング用の文字は、テキスト中で出現が検出されなかった文字が選択手段で選択される。なお、選択手段は、対象とするテキストの文字の出現状態を記憶する出現文字記憶手段を参照して、指定手段により指定されたマーキング用の文字より選択するようにしてもよい。

#### 【0007】

【実施例】図1は、この発明の一実施例に係る装置の概略構成図である。この図1において、入力部1は、テキスト処理を制御するための制御情報および処理対象とするテキストを前段の入力ファイル7から読み込み、対応する情報を後段のテキスト記憶部2や出現文字記憶部5に記憶させる機能を果たすものである。テキスト記憶部2は、入力部1を介して入力ファイル7から読み込まれた処理対象のテキストおよび制御情報を記憶する機能を果たすものである。テキスト検出処理部3は、テキスト記憶部2に記憶されているテキストから処理単位としての1文ずつを切り出し、この切り出した1文を構成

する文字列に対して前記制御情報で指定されている処理を施して、この処理結果を後段の処理結果記憶部4に記憶させる機能を果たすものである。また、その際に、処理対象のテキストにおけるある所定の文字の出現の有無について調査を施し、その結果を出現文字記憶部5に記憶させる機能も果たすようにされる。ここで対象のテキストに対する処理の一つとして、テキストを構成する文章中の誤字・脱字の存在を検出することがある。例えば、[1]ある所定の辞書に登録されていない単語、

[2]定義付けのない単語間の接続、[3]誤りを発生しやすい接続等を検出する技術については、例えば、特開平4-134408号公報に開示された方法を用いて実現することが可能である。この特開平4-134408号公報に開示されている方法によれば、処理対象とするテキストを構成する文章を単語や文節の単位に分割し、誤りを発生しやすいケースを検出するルールを適用することによって、該当する文章における誤字・脱字の候補を検出するようにされる。また、該当の文章を単語に分割する際には、単語の属性情報を記憶した辞書を参照することになるために、当該辞書に対応の[読み]が登録されているときには、[読み]が同じで[表記]が異なる単語を[表記不統一の候補]として検出することができる。処理結果記憶部4は、前段のテキスト検出処理部3によって検出された所定の情報を、処理対象のテキストに対応して記憶する機能を果たすものである。出現文字記憶部5は、対象とするテキストの文字の出現状態（例えば、テキストを構成する文章におけるある所定の文字の出現の有無のいかん）を記憶する機能を果たすものである。出力処理部6は、出現文字記憶部5における情報の内容を調査して対象テキスト中に出現していない文字[マーキング文字]を求め、この求めた[マーキング文字]を出力ファイル8に出力するとともに、対象テキストの所要の箇所にこの[マーキング文字]を挿入しながらこれも出力ファイル8に出力する機能を果たすものである。なお、この出力処理部6に含まれているものは、指定されたマーキング用の文字がテキスト中に出現したか否かを、前記出現文字記憶部5を参照して判断する判断部6A、前記出現文字記憶部5を参照してある所定の条件に合致した文字をマーキング用の文字として選択する選択部6B、上記のように選択されたマーキング用の文字を対象テキストの適所に挿入する挿入部6C、および、[マーキング用の文字]および[マーキング用の文字を挿入したテキスト]を出力する出力部6Dである。入力ファイル7は、機能的には入力部1の前段に配置されているファイルであり、処理対象としてのテキスト、および、その処理条件を含むようにされている。また、出力ファイル8は、機能的には出力処理部6の後段に配置されているファイルであり、前記[マーキング文字]それ自体、および、[要所にマーキング文字が挿入された対象テキスト]を保存するようにされている。

る。ここで、この図1におけるテキスト記憶部2は、特許請求の範囲の構成要件のテキスト記憶手段に対応している。テキスト検出処理部3は、検出処理手段に対応している。処理結果記憶部4および出力処理部6における挿入部6Cは、全体として挿入手段に対応している。そして、出現文字記憶部5、それに付属の指定部5A、出力処理部6における判断部6Aおよび選択部6Bは、全体として選択手段に対応している。また、出現文字記憶部5は出現文字記憶手段に対応しており、指定部5Aは指定手段に対応している。

【0008】以下、図2～図9を適宜参照しながら、前記図1に示された本実施例の動作について説明する。図2は、上記実施例の装置全体の動作を説明するためのフローチャートである。図3は、上記実施例の装置におけるマーキング文字の挿入動作を説明するためのフローチャートである。図4は、上記実施例の装置において処理できる入力ファイルの第1の例示図であり、その内容欄40が示されている。この図4において、前記の入力ファイル内容欄40は、ある所定の文字列（ここでは点線を表す文字列）によって、制御情報欄40Aおよび対象テキスト欄40Bに分割されている。なお、この図4においては、制御情報としては「誤り箇所を検出」が例示されている。また、対象テキストを構成する文章としては、「◎子プロセスがターミネートすると、親プロセスはsigcl dを受け終了する。sigcl dは親プロセスの修了をを知らせるのに役立つ。」なる文章が例示されている。図5は、上記実施例の装置における出現文字記憶部の初期状態の例示図である。この図5において、出現文字記憶部5の初期状態の内容の欄50には、対象とするテキストを構成する文章の文字欄50A、および、対応する文字の出現の有無欄50Bが設けられている。ここで、前記対象とするテキストを構成する文章の文字欄50Aには通常の文字は調査対象として含まれておらず、例えば、[※：@：\$：☆：★：○：●：◎：◇：◆：□：■]のような記号類だけが出現の有無の調査対象にされている。また、対応する文字の出現の有無欄50Bには、対応する記号毎に「出現ナシ：0」「出現アリ：1」のいずれかが記入される。なお、「出現アリ：1」については、その出現の回数には関係なく、1回であってもそれより多くの回数であっても「1」が記入されることになる。なお、この図5においては、出現文字記憶部5の初期状態の内容を表していることから、いずれの記号についても「出現ナシ：0」が記入されている。図6は、上記実施例の装置における出現文字記憶部の処理後の状態の例示図である。この図6においては、前記図5で例示された文章に「◎」なる記号が含まれていることから、この文章が切り出された後の出現文字記憶部5の状態の内容の欄60は次のようになる。即ち、対象とするテキストを構成する文章の文字欄60Aの中で、「◎」なる記号の箇所に対応

する文字の出現の有無欄60Bに「1」が立ち、他の箇所は「0」に留まるようにされる。図7は、上記実施例の装置における処理結果記憶部4の、ある1文を処理した後の状態の内容70の例示図である。いま、前記図4で例示された文章における「◎子プロセスがターミネートすると、親プロセスはsigcl dを受け終了する。：第1文」について、テキスト処理部3によって参照される辞書（図示されない）には「ターミネート」および「sigcl d」なる単語が登録されていないとすると、前者の単語「ターミネート」の該当文章における出現開始位置は「8」であり、その出現終了位置は「13」である（検出No1の欄70Aを参照）。（なお、ここでの出現開始位置等は、対象とする文章の最初の位置【ここでは、文字（記号）「◎」の存在位置】からカウントされる）。また、後者の単語「sigcl d」の該当文章における出現開始位置は「24」であり、その出現終了位置は「29」である（検出No2の欄70Bを参照）。ここでは、「ターミネート」および「sigcl d」なる単語のいずれも登録されていないことから、それらの誤り種別は「未登録語」にされている。図8は、上記実施例の装置における処理結果記憶部の、全文を処理した後の状態の例示図である。いま、前記図4で例示された文章における「sigcl dは親プロセスの修了をを知らせるのに役立つ。：第2文」について、テキスト処理部3によって参照される辞書には、前述されたように「sigcl d」なる単語が登録されておらず、また、単語「を」の連続「をを」についての定義付けがされていないとすると、それぞれに、図8における検出No3の欄80Cおよび検出No4の欄80Dのようにされる。即ち、単語「sigcl d」の該当文章における出現開始位置は「38」であり、その出現終了位置は「43」である（前記検出No3の欄80Cを参照）。そして、この「sigcl d」なる単語は登録されておらず、その誤り種別は「未登録語」にされている。また、単語「を」の連続「をを」については、前記検出No4の欄80Dに示されているように、その誤り種別は「接続未定義」にされる。なお、この図8における検出No1の欄80Aおよび検出No2の欄80Bは、それぞれに、内容的には図7における検出No1の欄70Aおよび検出No2の欄70Bと同等のものであり、ここで改めて説明することは省略する。そして、図9は、上記実施例の装置における第1の入力ファイルの内容（図4を参照）に対応する出力ファイルの内容90の例示図である。この図9には、図2および図3において後述されるステップS7（S701～S712）の作業が終了した結果としての内容が例示されている。この図9において、前記の出力ファイル内容欄90は、ある所定の文字列（ここでは点線を表す文字列）によって、制御情報欄90Aおよび（処理済み）

対象テキスト欄 90B に分割されている。なお、この図9においては、制御情報の一種として「マーキング文字：※」が示されている。また、（処理済み）対象テキストを構成する文章として、「◎子プロセスが※ターミネート※すると、親プロセスは※sigcl d※を受け終了する。※sigcl d※は親プロセスの修了を※を※知らせるのに役立つ。」が示されている。このような処理済みの対象テキストが得られることから、ワードプロセッサのオペレータは、探索パターンとして「※」なる文字（記号）を指定することにより、文章を作成する上で疑問が残る箇所を迅速かつ的確に探索・検出することができる。そして、ワードプロセッサのオペレータは、検出箇所について順次に確認し、必要に応じて修正を施す。このような確認・修正作業が終了したときには、前記「※」なる文字（記号）を「」なる空文字を表す記号に一括して置換することにより、「マーキング文字」としての「※」なる文字（記号）を確実に一括して削除することができる。ここで、前記「マーキング文字」としての「※」なる文字（記号）は、校正推敲対象のテキストとしての文章を構成するためには用いられておらず、この「マーキング文字」を削除することによって、前記された校正推敲対象のテキストの一部が誤って削除されることはない。

【0009】ここで図2を主として参照しながら、上記実施例の全体的な動作について説明していく。上記実施例の装置が起動されると、入力部1は、処理対象のテキストおよび対応の制御情報を入力ファイル7からテキスト記憶部2に記憶させる。その結果として、図4に例示されている対象テキスト、および、「誤り検出」を意味する制御情報が前記のテキスト記憶部2に記憶される

（S1）。テキスト検出処理部3は、前記テキスト記憶部2に記憶された対象テキストを構成する所定の1文を読み出す（S2）。ここでの文章の切り出しについては、文章中の句点「。」やピリオド「.」を手がかりとして、該当する文章中で多く現れるものを選択して所期の切り出しをする。この実施例においては、前記図4に例示されているように、句点「。」による切り出しがなされている。ここでは、前記された「◎子プロセスがターミネートすると、親プロセスはsigcl dを受け終了する。sigcl dは親プロセスの修了を知らせるのに役立つ。」なるテキストから文章を切り出す。そして、このような文章の切り出しのための句点「。」を探索する過程において、前記図5に例示されているような出現文字（記号）の有無が調査されて、その調査の結果が出現文字記憶部5に記憶される。現在の段階においては、前記された文章の中の「◎子プロセスがターミネートすると、親プロセスはsigcl dを受け終了する。：第1文」が記憶されることになる。テキスト検出処理部3は、上記のように切り出した1文について、テキスト記憶部2からの制御情報に従った処理（ここで

は、「誤り検出」なる制御情報に従った処理）を施す（S3）。なお、ここでの誤り検出の処理としては、例えば、前記された特開平4-134408号公報に開示されている方法を用いて実現することができる。即ち、既に触れてあることではあるが、「辞書に登録されていない単語」、「定義付けされていない単語間の接続」、「誤りやすい接続」等に該当することを検出することができる。次に、テキスト記憶部2に処理すべき文章が残留しているか否かを調べる（S4）。ここでは、処理すべき文章がまだ残留しているために先のステップS2に戻る。処理すべき文章が残留していないときには、次に続くステップS5に移行することになる。前述されたステップS2においては、「◎子プロセスがターミネートすると、親プロセスはsigcl dを受け終了する。sigcl dは親プロセスの修了を知らせるのに役立つ。」なるテキストから、「sigcl dは親プロセスの修了を知らせるのに役立つ。：第2文」が読み出される。ここでも、このような文章の切り出しのための句点「。」を探索する過程において、前記図6に例示されているような出現文字（記号）の有無が調査されて、その調査の結果が出現文字記憶部5に記憶される。なお、この段階においては、出現文字記憶部5における対応の文字（記号）は、いずれも前記の第2文に出現しておらず、出現文字記憶部5の記憶内容は、前記図6に示された状態に留まることになる。これに続くステップS3においては、前記の第2文について、テキスト記憶部2に記憶されている制御情報に従った処理（ここでも「誤り検出」なる制御情報に従った処理）が施される。次のステップS4においては、前記の第2文に続けて読み出すべき文章がテキスト記憶部2に残留しているか否かを調べる。この段階においては、第2文に続けて読み出すべき文章は存在せず、次に続くステップS5に移行する。このステップS5においては、出力処理部6の操作により、出現文字記憶部5の記憶内容（図6を参照）がその先頭から順次調べられて、「出現ナン」とされている最初の文字（記号）である「※」が、いわゆる「マーキング文字」として選択される。これに次いで、出力処理部6（内の出力部6D）からは、前記の「マーキング文字」である「※」が、出力ファイル8に対して出力される（S6）。これに続けて、出力処理部6（内の出力部6D）からは、テキスト記憶部2に記憶されている対象テキストが1文字ずつ出力ファイル8に対して出力される（S7）。この過程において、後述される図3のフローチャートに例示されているように、誤りの検出箇所に関する開始位置や終了位置に対応する文字を出力するのに先立ち、前記のように定められた「マーキング文字」を出力して、全体としての処理が終了する。

【0010】次に、主として図3のフローチャートを参照して、前記図2におけるステップS7に関する説明をする。この図3において用いられる各種の変数はそれぞれ

れに次のような意味を有するものである。即ち、  
cnt : カウンタ用の変数である。その初期値は 1 であり、転送しようとする文字の、対象テキストにおける位置を示すために用いられる。

maxCnt : 対象テキストの文字数を示す変数である。

No : 処理結果記憶部 4 に記憶された検出 No. を示す変数である。

MaxNo : 処理結果記憶部 4 に記憶された検出 No. の個数 (最大値) を示す変数である。

pos : 「マーキング文字」を挿入する位置を保持している変数である。処理結果記憶部 4 に記憶された検出 No. で指定される検出結果の開始位置または終了位置のいずれかを保持する。この pos で検出結果の開始位置または終了位置のいずれを指示するかは、後続の変数 flag で判定される。

flag : 前記 pos の値が検出結果の開始位置または終了位置のいずれを指示するかを判別する値を保持する変数である。この flag の値が 1 であれば開始位置であり、また、0 であれば終了位置である。

【0011】いま、図 3 においてステップ S6 が終了すると、前記された各変数の初期値がセットされる (S701)。ここでは、[cnt←1 : No←1 : flag←1 : pos←No. start (1 番目の検出結果の開始位置である 8) : maxCnt←64 (対象テキストの全文字数) : maxNo←4 (検出結果の個数)] のように初期値がセットされることになる。次に、変数 cnt の値が、別の変数 pos に等しいか否かを調べる (S702)。このステップ S702 においては、「マーキング文字」を挿入する位置であるかどうか判定される。ここでは、cnt = 1 であり、pos = 8 であって、両者は等しくないことから、ステップ S703 に移行して、前記 cnt の値に対応する文字 (記号) である [◎] が出力される。これに続くステップ S704 においては、前記 cnt の値が 1 だけ増大され、更に続くステップ S705 においては、前記 cnt の値が maxCnt の値を越えたか否かが判定される。このステップ S705 においてなされる判定は、対象テキストの全ての文字が出力したか否かについてである。ここでは cnt = 1, maxCnt 64 であり、その判定は否定であることから、元のステップ S702 に戻ることになる。cnt の値が 2, 3, 4, 5, 6, 7 であるときには、S702→S703→S704→S705→S702 の処理が繰り返され、[子], [プ], [ロ], [セ], [ス],

[が] のような文字が順次出力される。いま cnt の値が 8 であるとする、ステップ S702 における判定が肯定になり、対応のステップ S706 に移行する。このステップ S706 においては、検出結果の種別に応じた「マーキング文字」の出力がなされるが、ここで

例における「マーキング文字」は、検出結果の種別に係わらず、[※] であるものとされている。そこで、このステップ S706 においては、前記の [※] なる「マーキング文字」を出力してから、次のステップ S707 に移行する。このステップ S707 においては、flag = 1 であるか否か、即ち、pos の値が開始位置または終了位置のいずれに対応しているかが判定される。いまは flag = 1 であり、前記 pos の値が開始位置のものであることが示されているために、ステップ S709 に移行して、pos の値を、No. end + 1 (1 番目の検出結果の終了位置 : 13 + 1) である 14 にセットし、また、flag の値を 0 にセットする。これに続けてステップ S703 に移行し、ここでの cnt の値 (= 8) で示される位置の文字 [タ] を出力する。これに次いで、S704→S705→S702 の順に従う処理をする (なお、前記ステップ S704 において、cnt の値は 9 にセットされることになる)。cnt の値が 9, 10, 11, 12, 13 であるときには、S702→S703→S704→S705→S702 の処理が繰り返され、

[一], [ミ], [ネ], [ー], [ト] のような文字が順次出力される。いま cnt の値が 14 であるとする、ステップ S702 における判定が肯定になり、対応のステップ S706 に移行する。そして、[※] なる「マーキング文字」を出力してから、次のステップ S707 に移行する。このステップ S707 においては、flag の値が 0 であることからステップ S708 に移行し、No の値を 1 だけ増大させて (即ち、2 にして) からステップ S710 に移行する。このステップ S710 においては、No の値が maxNo の値 (= 4) を越えたか否かが判定される。即ち、所定の検出箇所に関するマーキング処理が全て終了したか否かが判定される。前記 No の値が maxNo の値を越えたと判定されたときには、ステップ S712 に移行して、pos の値を 0 にセットし、始めのステップ S702 における判定の結果が全て否定になるようにされる。いまは、前記 No の値が maxNo の値を越えておらず、ステップ S711 に移行して、pos の値を No. Start (2 番目の検出結果の開始位置) である 24 にセットし、これに続けてステップ S703 に移行する。このステップ S703 においては、cnt の値 (= 14) で示される位置に対応する文字 [す] が出力されて、S704→S705→S702 の処理がなされる (ここで、cnt の値は 15 にセットされることになる)。

【0012】以上説明したことは、cnt の値が 1 から 14 までの対象テキストの文字の出力をすること、および、処理結果記憶部 4 の検出 No1 における検出結果に対応するマーキング文字を挿入することである。この結果として、出力ファイル 8 には、[◎子プ



ロセスが※ターミネート※] なる文字列が出力される。以下、同様な作業をすることにより、処理結果記憶部4の 検出 No2, No3, No4 の開始位置、終了位置に対応する位置にマーキング文字が出力される。そして、ステップS705において最後に cnt の値が maxCnt の値を越えたときに、必要な処理が全て終了したことになる。

【0013】次に、図10～図16を適宜参照しながら、上記実施例に対して第2の内容例を有する入力ファイルが適用される場合について説明する。図10は、上記実施例に適用される第2例の入力ファイル内容欄100の例示図である。この図10において、前記の入力ファイル内容欄100は、ある所定の文字列（ここでは点線を表す文字列）によって、制御情報欄100A および対象テキスト欄100B に分割されている。なお、この図10においては、制御情報としては[誤り検出●] および[表記の不統一の検出◎] が例示されている。ここで、誤り箇所については[●] をもって囲み、表記不統一の箇所については[◎] をもって囲むようにされている。また、対象テキストを構成する文章としては、[◎子プロセスがターミネートすると、親プロセスはsigcldを受け終了する。sigcldは親プロセスの修了を知らせるのに役立つ。] なる文章が例示されている。出現文字記憶部5の内容の初期状態については、前記実施例における第1の例の入力ファイルの説明に係わる図5の場合と同様である。この出現文字記憶部5において予め指定されている文字（記号）とは異なるものが入力ファイル7において指定されたときには、当該文字（記号）を出現文字記憶部5に改めて登録するとともに、その出現の有無に関する情報として[0] を記入する。

【0014】ここで図11を主として参照しながら、上記実施例の第2の例の入力ファイルに係わる全体的な動作について説明していく。いま装置が起動されると、入力部1は、処理対象のテキストおよび対応の制御情報を入力ファイル7からテキスト記憶部2に記憶させる。その結果として、図10に例示されている対象テキスト、および、[誤り検出] および[表記の不統一の検出] を意味する制御情報が前記テキスト記憶部2に記憶される(S1)。テキスト検出処理部3は、前記テキスト記憶部2に記憶された対象テキストを構成する所定の1文を読み出す(S2)。その結果として、[◎子プロセスがターミネートすると、親プロセスはsigcldを受け終了する。] という1文が切り出される。また、対象とする文章の区切り文字を探索する過程において出現文字の有無を同時に調べていき、その結果を出現文字記憶部5に記憶させる。この場合、[◎] なる文字（記号）が出現していることから、前記の1文が切り出された後の出現文字記憶部5の内容は、前記された第1例の入力ファイルに係わる動作の場合と同様に、図6に示

されたようになる。テキスト検出処理部3は、上記のように切り出した1文について、テキスト記憶部2からの制御情報に従った処理（ここでは、[誤り検出] および[表記の不統一の検出] なる制御情報に従った処理）を施す(S3)。なお、ここでの誤り検出の処理としては、前記第1例の入力ファイルに係わる動作の場合において例示された方法を用いて実現することができる。また、表記の不統一の検出については、[読み] は同じであるが[表記] が異なるもの（同音異義語、送りがなのゆれ、漢字表記とかな表記との混在）を検出するようにされる。ここで、処理結果記憶部4の内容についてみると、誤り検出については、前記第1例の入力ファイルに係わる動作の場合と同様に、図7に示されているようになる。また、表記の不統一については図12に例示されたようになる。この図12は、上記第2第1例の入力ファイルに係わる動作の場合の処理結果記憶部の、ある1文を処理した後の表記に関する状態の例示図である。この図12において、見出し部121に含まれている欄は、読み欄121A、表記数欄121B、表記への番地欄121Cであり、例えばその第12A段には、[読み:こ]、[表記数:1] および[表記への番地:1] が記入されている。また、表記部122に含まれている欄は、番地欄122A、次の表記の番地欄122B、開始位置欄122C、終了位置欄122Dであり、例えばその第12E段には、[番地:1] [次の表記の番地:0]、[開始位置:2] および[終了位置:2] が記入されている。即ち、この図12においては、こ（子）、おや（親）、うけ（受け）、しゅうりょう（終了）という単語の[読み] および[出現位置] が、見出し部121および表記部122に分けて格納されている。ここでの[出現位置] は、前記第1例の入力ファイルに係わる動作における図7および図8の場合と同様に、ある所定の単語の[開始位置] および[終了位置] から定まるものである。この図12における[表記数:121B] は、読みが同じで表記が異なる単語の出現回数を示している。この値が例えば2であるときには、表記に不統一な箇所があったものと判断されることになる。ステップS3に続き、テキスト記憶部2には、処理すべき文章が残留しているか否かを調べる(S4)。処理すべき文章が残留しているときには、ステップS5に移行する。これに対して、処理すべき文章が残留していないときには、始めのステップS2に戻ることになる。テキスト検出処理部3は、テキスト記憶部2に記憶されている対象テキストから、次に続く1文を読み出す(S2)。その結果として、[sigcldは親プロセスの修了を知らせるのに役立つ。] なる1文が読み出される。また、文章の区切り文字（記号）を探索する過程において、出現文字の有無についても同時に調べ、その結果を出現文字記憶部5に記憶させる。この場合においては、該出現文字記憶部5に格納されている

文字（記号）はいずれも出現しておらず、このような処理が施された後の出現文字記憶部5の内容は、前記第1例の入力ファイルに係わる動作における図6と同様のものである。テキスト検出処理部3は、前記のように切り出された1文について、テキスト記憶部2における制御情報に従った処理を施す（S3）。このときの処理結果記憶部4の内容について、[誤り検出]については前記第1例の入力ファイルに係わる動作における図8と同様になり、また、[表記の不統一]については図13に示されているようになる。この図13は、上記第2例の入力ファイルに係わる動作における処理結果記憶部の、全文を処理した後の表記に関する状態の例示図である。この図13において、見出し部131に含まれている欄は、読み欄131A、表記数欄131B、表記への番地欄131Cであり、例えばその第13B段には、[読み：おや]、[表記数：1]および[表記への番地：2]が記入されている。また、表記部132に含まれている欄は、番地欄132A、次の表記の番地欄132B、開始位置欄132C、終了位置欄132Dであり、例えばその第13G段には、[番地：1] [次の表記の番地：0]、[開始位置：2]および[終了位置：2]が記入されている。ここで、おや（親）という単語についてみると、同じ表記が既に登録されており、改めてこれを登録することはない。ただし、しゅうりょう（修了）という単語については、異なる表記である[終了]が登録されており、ここでは表記数を[2]とし、表記部132の適所に[修了]を登録する。これに続けて、切り出すべき文字列がテキスト記憶部2にまだ残っているか否かを調べる。ここでは、切り出すべき文字列が残っておらず、このために、次のステップS5に移行して、使用すべき[マーキング文字]を決定する。このとき、出力処理部6においては、テキスト記憶部2に記憶されている対象テキスト中での[マーキング文字]の出現の有無について、出現文字記憶部5の内容を先頭から順次に調べて確認する。ここで[出現アリ：1]のものがあるときには、前記出現文字記憶部5における[出現ナシ：0]のものから適当なものを選択する。この場合において、[◎]なる文字（記号）が対象テキスト内で既に出現していることから、例えば、これに続く[◇]なる文字（記号）を[マーキング文字]として選択する。この結果として、図14に示されているような[検出の種別]と[マーキング文字]との対応テーブル140が形成される。この図14は、上記第2例の入力ファイルに係わる動作における検出種別とマーキング文字との対応例示図である。この図14における対応テーブル140には、[誤り種別欄：141]、[マーキング文字欄：142]および[指定されたマーキング文字欄：143]が含まれている。この図14においては、第140A段および第140B段に示されているように、[未登録語]および[接続未定義]につい

てはそれぞれに[指定されたマーキング文字]である[●]を[マーキング文字]として用いる。これに対し、[表記の不統一]については[指定されたマーキング文字]である[◎]ではなく、別の[◇]を[マーキング文字]として用いる。図11のフローチャートにおいては、前記のステップS5に続けて、異なる処理結果を一つにまとめる（S5A）。この場合、前記された図8における誤り検出結果の状態、および、図13における表記の不統一に関する検出結果の状態について、その開始位置の小さい順に並ぶようにマージして、図15に示されているような処理結果のテーブルにされる。この図15は、上記第2例の入力ファイルに係わる動作における処理結果記憶部の、全文を処理した後の状態のテーブル状の例示図である。この図15における処理結果テーブル150には、[検出No欄：151]、[開始位置欄：152]、[終了位置欄：153]、[誤り候補欄：154]および[誤り種別欄：155]が含まれている。このような処理結果テーブル150を用意することにより、対象テキストを構成する文章の適所に、適切な[マーキング文字]を確実に挿入することができる。図11のフローチャートにおける前記のステップS5Aに続けて、出力処理部6は、図15における対応テーブル150を参照しながら、出力ファイル8に対して適切な[マーキング文字]を出力させる（S6）。ここで、実際に用いられる[マーキング文字]が[指定されたマーキング文字]とは異なるときには、その変更が認識されるような出力をさせることになる。ここでの例においては、誤り種別が[表記の不統一]であるときに、実際に用いられる[マーキング文字]は[指定されたマーキング文字]である[◎]から[◇]に変更されている。図16は、上記第2例の入力ファイルに係わる動作における出力ファイルの内容160の例示図である。この図16において、前記の出力ファイル内容欄160は、ある所定の文字列（ここでは点線を表す文字列）によって、制御情報欄160Aおよび（処理済み）対象テキスト欄160Bに分割されている。なお、この図16においては、前者の制御情報である[誤り検出]に対する[マーキング文字]として[●]が示されており、また、後者の制御情報である[表記の不統一の検出]に対する[マーキング文字]として[◎]に代わる[◇]が示されている。ステップS6の結果は、図16の160Bに示されているように、[◎子プロセスが●ターミネート●すると、親プロセスは●sigcl d●を受け◇終了◇する。●sigcl d●は親プロセスの◇修了◇を●を●知らせるのに役立つ。]となる。図11のフローチャートにおいては、前記のステップS6に続けて、出力処理部6は、テキスト記憶部2からの対象テキストを、1文字ずつ出力ファイル8に対して転送する（S7）。この転送の過程において、前記第1例の入力ファイルに係わる動作に関する図3のフローチャ

ートを参照すると、誤り候補の検出箇所の開始位置や終了位置に対応する文字を出力するのに先だって対応の「マーキング文字」を出力するようにされる。なお、ここでの第2例の入力ファイルに係わる動作に関する場合には、前記図3のフローチャートのステップS706においては、図14における対応テーブル140を参照しながら、誤りの種別に応じた「マーキング文字」を出力するようにされる。前述されたように、図16は、前記図11のフローチャートのステップS7が終了したときに、結果として出力された内容の例示図である。この図16の場合においては、[●]および[◇]が「マーキング文字」として用いられているが、そのいずれも対象テキストには含まれていない。このために、使用されるワードプロセッサの探索パターンとして、前記[●]および[◇]を指定することにより、該当の検出箇所を迅速かつ的確に探索することができる。そして、検出された箇所については逐次を確認し、必要に応じて修正を施す。そして、必要な確認・修正が終了したときには、これらの[●]および[◇]を□（空文字）で一括置換することで、対象テキスト内の「マーキング文字」だけを確実に削除することができる。前述された第1例の入力ファイルの場合と同様に、対象テキスト内には「マーキング文字」に一致した文字（記号）は用いられておらず、前記「マーキング文字」の削除操作において、校正推敲対象のテキストの一部が誤って削除されることが確実に防止される。

【0015】図17は、上記実施例において適用される第3例の入力ファイルの内容例示図である。この図17において、前記の入力ファイル内容欄170は、ある所定の文字列（ここでは点線を表す文字列）によって、制御情報欄170Aおよび対象テキスト欄170Bに分割されている。なお、この図17においては、制御情報としては「誤り検出：●」および「表記の不統一の検出：◎：KWIC」が例示されている。また、対象テキストを構成する文章としては、他の入力ファイル内容例の場合と同様に、[◎子プロセスがターミネートすると、親プロセスはsigcldを受け終了する。sigcldは親プロセスの終了を知らせるのに役立つ。]なる文章が例示されている。ここで、KWICとは、[Key Word In Context]を略したものであり、キーとなる文字列を、その前後の文字列とともに表現したものである。ここでの表記が不統一な表現としては、複数ページにまたがって離れて出現することがあり、対象テキストのページを繰り返しながら確認していくことは困難である。そこで、例えば、[読み]が同じで[表記]が異なるものをキーとしてKWIC表示することにより、表記の不統一な箇所を一覧で確認できるようになる。この第3例の入力ファイルの場合には、その動作は前記第2例の入力ファイルの場合と同様であり、指定された「マーキング文字」が挿入された出力

は、前記図16と同様になる。また、出力処理部6は、処理結果記憶部4を参照しながら、図18に示されているようなKWIC表示の出力をする。この図18は、上記第3例の入力ファイルの場合におけるKWICに関する出力ファイルの内容180の例示図である。この図18におけるKWIC部を参照することにより、その前後関係を見ることによって表記に不統一があることが容易に確認される。この図18においては、[しゅうりょう]なる[読み]180Aが調査の対象にされており、[終了]および[修了]なる表記180Bが不統一のものとされている。ここでは、前記された「表記」の中の[終了]が正しいものとされ、これに合わせた文章の修正作業がなされることになる。なお、この第3例の入力ファイルに係わる動作の場合にはKWICなる手法が用いられたけれども、これ以外に、単語の頻度リストの作成依頼とそのリストの作成；複数の文字列の探索置換の依頼と置換のログの作成；というような処理も実施することが可能である。

#### 【0016】

【発明の効果】以上説明されたように、この発明によれば、処理対象のテキストに関して修正を要する箇所については、当該処理対象のテキストには出現していない文字（記号）を「マーキング文字」として選択し、挿入することによって該当の箇所が明示されており、この該当の検出箇所については、利用者が日常的に使用している通常のワードプロセッサに本来的に付属している基本的な機能の一つである探索機能を用いることにより、迅速かつ的確な検出および所要の修正処理を施すことができる。

#### 【図面の簡単な説明】

【図1】この発明の一実施例に係る装置の概略構成図である。

【図2】上記実施例の装置全体の動作を説明するためのフローチャートである。

【図3】上記実施例の装置におけるマーキング文字の挿入動作を説明するためのフローチャートである。

【図4】上記実施例の装置における入力ファイルの第1の例示図である。

【図5】上記実施例の装置における出現文字記憶部の初期状態の例示図である。

【図6】上記実施例の装置における出現文字記憶部の処理後の状態の例示図である。

【図7】上記実施例の装置における処理結果記憶部の、ある1文を処理した後の状態の例示図である。

【図8】上記実施例の装置における処理結果記憶部の、全文を処理した後の状態の例示図である。

【図9】上記実施例の装置における第1例の入力ファイルの内容に対応する出力ファイルの内容の例示図である。

【図10】上記実施例の装置に適用される第2の入力フ

ファイル内容の例示図である。

【図11】上記実施例の装置に適用される第2の入力ファイルに係わる全体的な動作を説明するためのフローチャートである。

【図12】上記第2例の入力ファイルに係わる動作における処理結果記憶部の、ある1文を処理した後の表記に関する状態の例示図である。

【図13】上記第2例の入力ファイルに係わる動作における処理結果記憶部の、全文を処理した後の表記に関する状態の例示図である。

【図14】上記第2例の入力ファイルに係わる動作における検出種別とマーキング文字との対応例示図である。

【図15】上記第2例の入力ファイルに係わる動作における処理結果記憶部の、全文を処理した後の状態のテー

ブル状の例示図である。

【図16】上記第2例の入力ファイルに係わる動作における出力ファイルの内容の例示図である。

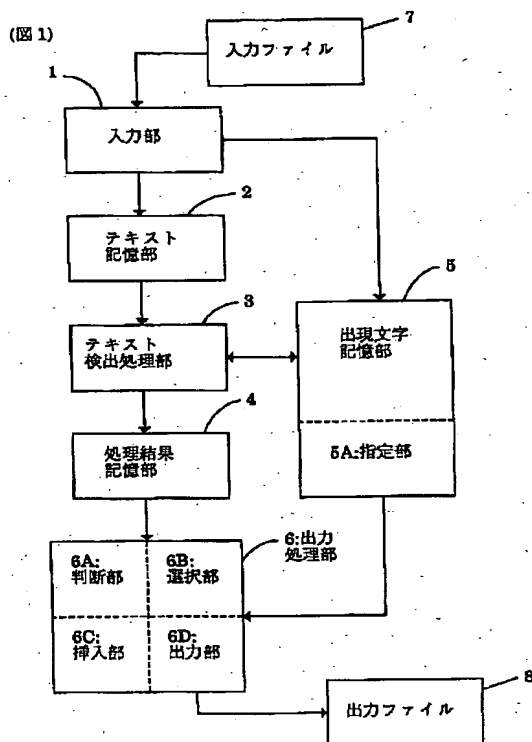
【図17】上記実施例において適用される第3の入力ファイルの内容例示図である。

【図18】上記第3の入力ファイルの場合における K W I C に関する出力ファイルの内容の例示図である。

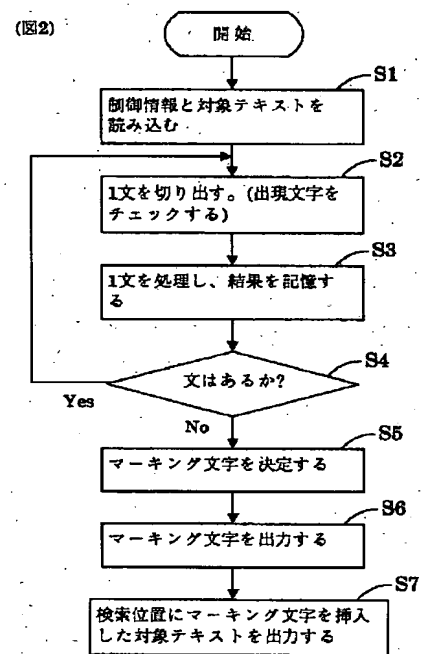
【符号の説明】

1—入力部；2—テキスト記憶部；3—テキスト検出処理部；4—処理結果記憶部；5—出現文字記憶部；5A—指定部；6—出力処理部；6A—判断部；6B—選択部；6C—挿入部；6D—出力部。

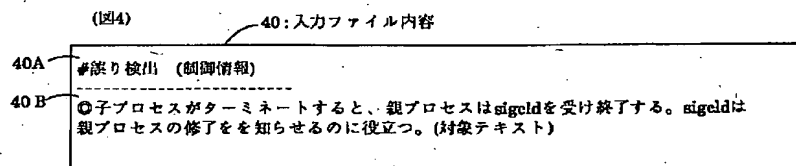
【図1】



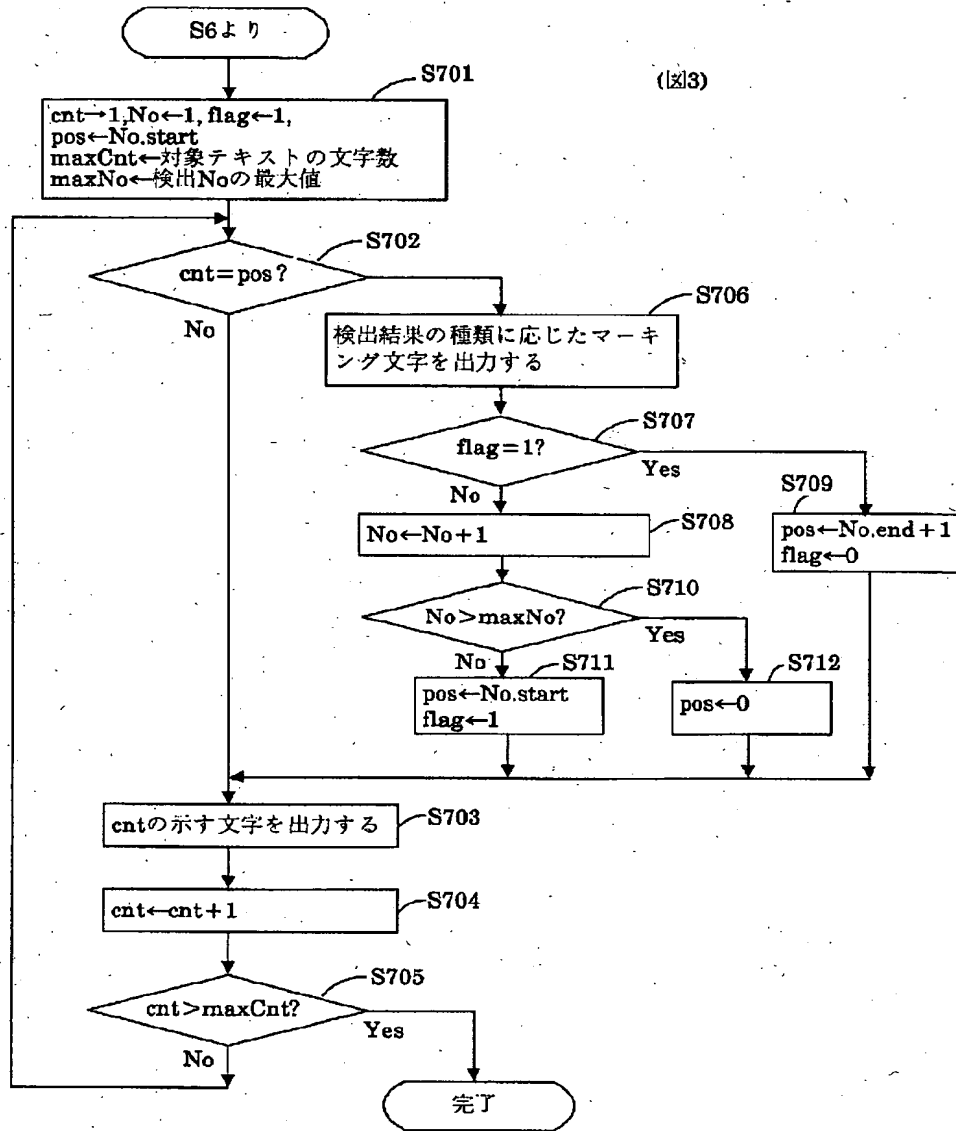
【図2】



【図4】



【図3】



【図5】

(図5) 50: 手段5の初期状態の内容

50A	文字	※	@	§	☆	★	○	●	◎	◇	◆	□	■
50B	出現の有無	0	0	0	0	0	0	0	0	0	0	0	0

【図6】

(図6) 60: 手段5の処理状態の内容

60A	文字	※	@	\$	☆	★	○	●	◎	◇	◆	□	■
60B	出現の有無	0	0	0	0	0	0	0	1	0	0	0	0

【図7】

(図7) 70: 手段4の処理後の内容

	検出 No	開始 位置	終了 位置	誤り候補	誤り種別
70A	1	8	13	ターミネート	未登録語
70B	2	24	29	sigcld	未登録語

【図8】

(図8) 80: 手段4の最終処理後の内容

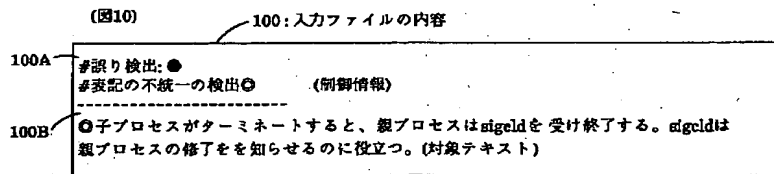
	検出 No	開始 位置	終了 位置	誤り候補	誤り種別
80A	1	8	13	ターミネート	未登録語
80B	2	24	29	sigcld	未登録語
80C	3	38	43	sigcld	未登録語
80D	4	54	54	を	接続未定義

【図9】

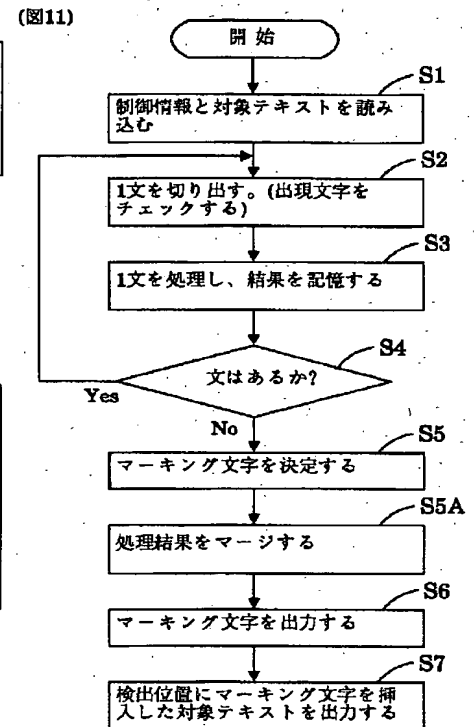
(図9) 90: 出力ファイルの内容

90A	#マーキング文字:※ (制御情報)
90B	○子プロセスが※ターミネート※すると、親プロセスは※sigcld※を受け終了する。※sigcld※は親プロセスの終了を※を※知らせるのに役立つ。(対象テキスト)

【図10】



【図11】



【図14】

(図14)

	141 誤り種別	142 マーキング文字	143 指定された マーキング文字
140A	未登録語	●	●
140B	接続未定義	●	●
140C	表記の不統一	◇	◎

140: 対応テーブル

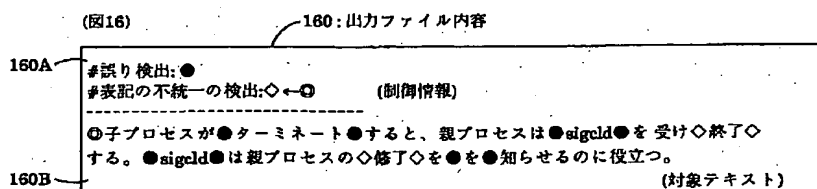
【図12】

(図12)

	121A 読み	121B 表記数	121C 表記への 番地	122A 番地	122B 次の表記の 番地	122C 開始位置	122D 終了位置	
12A	こ	1	1	1	0	2	2	12E
12B	おや	1	2	2	0	18	18	12F
12C	うけ	1	3	3	0	31	32	12G
12D	しゅうりょう	1	4	4	0	33	34	12H

121: 見出し部      122: 表記部

【図16】



【図13】

(図13)			131A	131B	131C	132A	132B	132C	132D	
	読み	表記数	表記への 番地	番地	次の表記の 番地	開始位置	終了位置			
13A	こ	1	1	1	0	2	2	13G		
13B	おや	1	2	2	0	18	18	13H		
13C	うけ	1	3	3	0	31	32	13J		
13D	しゅうりょう	2	4	4	5	33	34	13K		
13E	しらせ	1	6	5	0	51	52	13L		
13F	やくだ	1	7	6	0	55	56	13M		
				7	0	61	62	13N		

131: 見出し部

132: 表記部

【図15】

(図15)	151	152	153	154	155
	検出 No	開始 位置	終了 位置	誤り候補	誤り種別
150A	1	8	13	ターミネート	未登録語
150B	2	24	29	sigcld	未登録語
150C	3	33	34	終了	表記の不統一
150D	4	38	43	sigcld	未登録語
150E	6	51	52	修了	表記の不統一
150F	5	54	54	を	接続未定義

150: 処理結果テーブル

【図17】

(図17)	170: 入力ファイル内容
170A	#誤り検出: ● #表記の不統一の検出: ○: KWIC (制御情報)
170B	○子プロセスがターミネートすると、親プロセスはsigcldを受け終了する。sigcldは 親プロセスの修了を知らせるのに役立つ。 (対象テキスト)

【図18】

(図18)	180: 出力ファイル内容
180A	読み: しゅうりょう
180B	( sigcldを受け 終了 する。 プロセスの 修了 を知らせるの